Aggregation Principle in Stochastic Optimization

Progressive Hedging Algorithm

ENSTA, Paris, Summer 2014.

Dealing with Uncertainty

in Decision Making Models 2.

Roger J-B Wets

Stochastic Optimization

1. **Recourse model**: $\min Ef(x) = \mathbb{E}\{f(\xi, x)\}$

$$f(\xi, x) = \begin{cases} f_{01}(x) + Q(\xi, x) & \text{if } x \in C_1 \\ \infty & \text{otherwise} \end{cases}$$

$$Q(\xi, x) = \inf_y \left\{ f_{02}(\xi, y) \,\middle|\, y \in C_2(\xi, x) \right\}$$

2. $\sim$ **Pricing model**: $\min Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi}, x(\boldsymbol{\xi}))\}$

$$x(\xi) = \big(x_0(\xi), x_1(\xi), \dots, x_T(\xi)\big), \;\; x_t(\xi) \sim x_t(\overset{\to t}{\xi})$$

information field (filtration): $\left\{\mathcal{A}_0 = \{\emptyset, \Xi\}, \mathcal{A}_1, \dots, \mathcal{A}_T = \mathcal{A}\right\}$

$x_t : \Xi \to \mathbb{R}^{n_t}, \mathcal{A}_t$-measurable

Unit Commitment - Short version

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \qquad \text{with}$$

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

$$\text{Minimize} \sum_{k\in K}\sum_{j\in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \qquad \text{with}$$

J generating units

$$\sum_{j\in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j\in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

$$\text{Minimize} \sum_{k\in K} \sum_{j\in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \qquad \text{with}$$

K time periods *J generating units*

$$\sum_{j\in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j\in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \qquad \text{with}$$

K time periods *J generating units*

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost*

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \qquad \text{with}$$

K time periods *J generating units*

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost* *shutdown cost*

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods *J generating units*

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods.
The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost* *shutdown cost*

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods *J generating units*

demand

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost* *shutdown cost*

$$\text{Minimize} \sum_{k\in K} \sum_{j\in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods *J generating units*

power output *demand*

$$\sum_{j\in J} p_j(k) = D(k), \quad \forall k \in K$$

$$\sum_{j\in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods.
The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost   startup cost   shutdown cost

$$\text{Minimize} \sum_{k\in K}\sum_{j\in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods   J generating units

power output   demand

$$\sum_{j\in J} p_j(k) = D(k), \quad \forall k \in K$$

max power output

$$\sum_{j\in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost* *shutdown cost*

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods *J generating units*

power output *demand*

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

max power output *spinning reserve*

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods. The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost* *shutdown cost*

$$\text{Minimize} \sum_{k\in K}\sum_{j\in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods *J generating units*

power output *demand*

$$\sum_{j\in J} p_j(k) = D(k), \quad \forall k \in K$$

max power output *spinning reserve*

$$\sum_{j\in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods.
The specific nature of Π is model-dependent.

Unit Commitment - Short version

production cost *startup cost* *shutdown cost*

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods *J generating units*

demand

power output $$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

spinning reserve

max power output $$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods.
The specific nature of Π is model-dependent.

"Stochastic Version"

Unit Commitment - Short version

min. expectation
(actually: risk measure)
with penalties

production cost   startup cost   shutdown cost

$$\text{Minimize} \sum_{k\in K}\sum_{j\in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods   J generating units

power output   demand

$$\sum_{j\in J} p_j(k) = D(k), \quad \forall k \in K$$

adjust node
balance eq'ns

max power output   spinning reserve

$$\sum_{j\in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \ \forall k \in K$$

Π region of feasible production, all generating units, all time periods.
The specific nature of Π is model-dependent.

"Stochastic Version"

Interchanging: $\mathbb{E}$ & $\min$

Evident: with $E = \{x : \Xi \to \mathbb{R}^N \,|\,$ measurable, ... $\}$

$\min \mathbb{E}\left\{ f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big) \,\big|\, x \in E \right\} = \mathbb{E}\left\{ \min f(\boldsymbol{\xi}, x) \,\big|\, x \in \mathbb{R}^N \right\}$

when $\exists\, x(\,\cdot\,) \in E$ such that P-$a.s.$ $x(\xi) \in \operatorname{argmin} f(\xi, \,\cdot\,)$

x is measurable, ...

But our problem is: $\min \mathbb{E}\{f(\boldsymbol{\xi}, x)\}$, equivalently,

$$\min E f(x) = \mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big)\}$$

$$\text{such that} \quad x(\xi) = \mathbb{E}\{x(\boldsymbol{\xi})\} \; P\text{–}a.s.$$

x can not depend on ‘anticipated’ (future) information

Here-&-Now vs. Wait-&-See

- Basic Process: decision → observation → decision
  $x^1 \rightarrow \xi \rightarrow x^2_\xi$
- Here-&-Now problem! x^1 not all contingencies can be "protected" by available instruments / tools (in stage 1)
- Wait-&-See problem:
  instruments are available to cover all contingencies
  choose (x^1_ξ, x^2_ξ) after observing random event

Stochastic Optimization: Fundamental Theorem

A here-&-now problem can be transformed in a wait-&-see problem by introducing the

appropriate `contingencies' costs (price of nonanticipativity)

Price of Nonanticipativity

$$\begin{aligned} \min \ & \mathbb{E}\left\{ f(\xi, x^1, x_\xi^2) \right\} \\ & x^1 \in C^1 \subset \mathbb{R}^n, \\ & x_\xi^2 \in C^2(\xi, x^1), \forall \xi. \end{aligned}$$

Price of Nonanticipativity

$$\min \ \mathbb{E}\left\{ f(\xi, x^1, x^2_\xi) \right\}$$
$$x^1 \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1), \forall \xi.$$

Explicit non-anticipativity

$$\min \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) \right\}$$
$$x^1_\xi \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1_\xi), \forall \xi.$$

Price of Nonanticipativity

$$\min \ \mathbb{E}\left\{ f(\xi, x^1, x^2_\xi) \right\}$$
$$x^1 \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1), \forall \xi.$$

Explicit non-anticipativity

$$\min \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) \right\}$$
$$x^1_\xi \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1_\xi), \forall \xi.$$
$$x^1_\xi = \mathbb{E}\left\{ x^1_\xi \right\} \quad \forall \xi$$

Price of Nonanticipativity

$$\min \ \mathbb{E}\left\{ f(\xi, x^1, x^2_\xi) \right\}$$
$$x^1 \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1), \forall \xi.$$

Explicit non-anticipativity

$$\min \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) \right\}$$
$$x^1_\xi \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1_\xi), \forall \xi.$$

$$x^1_\xi = \mathbb{E}\left\{ x^1_\xi \right\} \quad \forall \xi$$

$w_\xi \perp$ subspace of constant fcns

$$\Rightarrow \mathbb{E}\left\{ w_\xi \right\} = 0$$

Price of Nonanticipativity

$$\min \ \mathbb{E}\left\{ f(\xi, x^1, x^2_\xi) \right\}$$
$$x^1 \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1), \ \forall \xi.$$

Explicit non-anticipativity

$$\min \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) \right\}$$
$$x^1_\xi \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1_\xi), \ \forall \xi.$$

$$x^1_\xi = \mathbb{E}\left\{ x^1_\xi \right\} \quad \forall \xi$$

$w_\xi \perp$ subspace of constant fcns

$$\Rightarrow \mathbb{E}\left\{ w_\xi \right\} = 0$$

$$\min \ \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) - \langle w_\xi, x^1_\xi \rangle + \langle w_\xi, \mathbb{E}\{x^1_\xi\} \rangle \right\}$$

such that $x^1_\xi \in C_1, \quad x^2_\xi \in C_2(\xi, x^1_\xi)$

Price of Nonanticipativity

$$\min \ \mathbb{E}\left\{ f(\xi, x^1, x^2_\xi) \right\}$$
$$x^1 \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1), \forall \xi.$$

Explicit non-anticipativity

$$\min \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) \right\}$$
$$x^1_\xi \in C^1 \subset \mathbb{R}^n,$$
$$x^2_\xi \in C^2(\xi, x^1_\xi), \forall \xi.$$

$$x^1_\xi = \mathbb{E}\left\{ x^1_\xi \right\} \quad \forall \xi$$

$w_\xi \perp$ subspace of constant fcns

$$\Rightarrow \mathbb{E}\left\{ w_\xi \right\} = 0$$

$$\min \ \mathbb{E}\left\{ f(\boldsymbol{\xi}, x^1_\xi, x^2_\xi) - \langle w_\xi, x^1_\xi \rangle + \langle w_\xi, \mathbb{E}\{x^1_\xi\} \rangle \right\}$$

such that $x^1_\xi \in C_1, \quad x^2_\xi \in C_2(\xi, x^1_\xi)$

Nonanticipativity

Recall $\quad \min Ef(x) = \mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big)\}$ such that $x(\xi) = \mathbb{E}\{x(\boldsymbol{\xi})\}$ P–$a.s.$

Nonanticipativity <u>constraints</u>:

$\mathcal{N}_a = \big\{x : \Xi \to \mathbb{R}^n\big\} \subset$ linear subspace of constant fcns

$\Longrightarrow \exists\, w : \Xi \to \mathbb{R}$ "multipliers" $\perp \mathcal{N}_a$ $(\Rightarrow \mathbb{E}\{w(\boldsymbol{\xi})\} = 0)$ such that

$$x^* \in \operatorname{argmin} Ef \quad \Longrightarrow \quad x^* \in \operatorname{argmin} \big\{\mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big) + \langle w(\boldsymbol{\xi}), (x(\boldsymbol{\xi}) - \mathbb{E}\{x(\boldsymbol{\xi})\})\rangle\}$$

$$\Longrightarrow \quad x^* \in \operatorname{argmin} \big\{\mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big) + \langle w(\boldsymbol{\xi}), x(\boldsymbol{\xi})\rangle\}\big\}$$

$$P\text{–}a.s. \quad \Longrightarrow \quad x^* \in \operatorname*{argmin}_{x \in E}\{f\big(\xi, x\big) + \langle w(\xi), x\rangle\}\big\}, \quad \xi \in \Xi$$

$w(.)$: contingencies equilibrium prices, $\sim$ 'insurance' prices

Dynamic Information Process

So far, x restricted to $\{\emptyset, \Xi\}$-measurable, i.e., constant on Ξ

Generally, as $t \nearrow T$ (possibly ∞) additional information is acquired

$$\mathcal{A}_0 = \{\emptyset, \Xi\} \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_T = \mathcal{A}, \text{ a filtration}$$

with x_t decision @ time t depend on available information, i.e. $\mathcal{A}_t$-measurable

<u>Reformulation</u>

Let $x(\xi) = \big(x_0(\xi), x_1(\xi), \ldots, x_T(\xi)\big) : \Xi \to \mathbb{R}^N, \; N = \sum_{t=0}^T n_t$

$$\mathcal{N}_a = \big\{ x \in E \,\big|\, x_t \; \mathcal{A}_t\text{-measurable}, \; t = 0, \ldots T \big\}$$

find $x \in \mathcal{N}_a$ such that $Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi}, x(\boldsymbol{\xi}))\}$ is minimized

Nonanticipativity <u>constraints</u>: $x \in \mathcal{N}_a$ (linear subspace)

Adjusted Here-&-Now

$\min\ \mathbb{E}\left\{f(\boldsymbol{\xi},x^1,x^2_\xi)\right\}$ such that $x^1 \in C^1 \subset \mathbb{R}^n,\ x^2_\xi \in C^2(\xi,x^1),\forall\xi$

x^1 must be $\mathcal{G}$-measurable, $\mathcal{G} = \sigma\text{-}\{\varnothing,\Xi\}$

x^2 is $\mathcal{A}$-measurable, $\mathcal{A} \supset \mathcal{G}$,

in general, interchange $\mathbb{E}$ & ∂ is not valid

required: $\forall\ \xi, x^1 \in C^1, C^2(\xi,x^1) \neq \varnothing$ $\mathcal{G}$-measurability of constraints

Now, suppose w_ξ are the (optimal) non-anticipativity multipliers (prices)

$\min\ \mathbb{E}\left\{f(\boldsymbol{\xi},x^1_\xi,x^2_\xi) - \langle w_\xi,x^1_\xi\rangle + \langle w_\xi,\mathbb{E}\{x^1_\xi\}\rangle\right\}$

such that $x^1_\xi \in C^1 \subset \mathbb{R}^n,\ x^2_\xi \in C^2(\xi,x^1_\xi),\forall\xi$

Interchange is now O.K. , $\mathbb{E}\left\{\langle w_\xi,\mathbb{E}\{x^1_\xi\}\rangle\right\} = \langle\mathbb{E}\{w_\xi\},\mathbb{E}\{x^1_\xi\}\rangle = 0$, yields

$\forall\ \xi$, solve: $\min\ f(\boldsymbol{\xi},x^1,x^2) - \langle w_\xi,x^1\rangle$ s.t. $x^1 \in C^1,\ x^2 \in C^2(\xi,x^1)$

a collection of deterministic optimization problems in $\mathbb{R}^{n_1+n_2}$

Finding w_ξ

Progressive Hedging Algorithm

0. $w^0(\cdot)$ such that $\mathbb{E}\{w^0(\xi)\} = 0,\ \nu = 0$. Pick $\rho > 0$

1. for all ξ :

$$(x_\xi^{1,\nu}, x_\xi^{2,\nu}) \in \arg\min f(\xi; x^1, x^2) - \langle w_\xi^\nu, x^1 \rangle$$

$$x^1 \in C^1 \subset \mathbb{R}^{n_1},\ x^2 \in C^2(\xi, x^1) \subset \mathbb{R}^{n_2}$$

2. $\bar{x}^{1,\nu} = \mathbb{E}\{x_\xi^{1,\nu}\}$. Stop if $\left| x_\xi^{1,\nu} - \bar{x}^{1,\nu} \right| = 0$ (approx.)

otherwise $w_\xi^{\nu+1} = w_\xi^\nu + \rho\left[x_\xi^{1,\nu} - \bar{x}^{1,\nu}\right]$, return to 1. with $\nu = \nu + 1$

Progressive Hedging Algorithm

0. $w^0(\cdot)$ such that $\mathbb{E}\{w^0(\xi)\} = 0,\ \nu = 0.$ Pick $\rho > 0$

1. for all ξ :

$$(x_\xi^{1,\nu}, x_\xi^{2,\nu}) \in \arg\min f(\xi; x^1, x^2) - \langle w_\xi^\nu, x^1 \rangle$$

$$x^1 \in C^1 \subset \mathbb{R}^{n_1},\ x^2 \in C^2(\xi, x^1) \subset \mathbb{R}^{n_2}$$

2. $\bar{x}^{1,\nu} = \mathbb{E}\{x_\xi^{1,\nu}\}$. Stop if $\left|x_\xi^{1,\nu} - \bar{x}^{1,\nu}\right| = 0$ (approx.)

otherwise $w_\xi^{\nu+1} = w_\xi^\nu + \rho\left[x_\xi^{1,\nu} - \bar{x}^{1,\nu}\right]$, return to 1. with $\nu = \nu + 1$

Convergence: add a proximal term

$$f(\xi; x^1, x^2) - \langle w_\xi^\nu, x^1 \rangle - \frac{\rho}{2}\left|x^1 - \bar{x}^{1,\nu}\right|^2$$

linear rate in $(\mathrm{x}^{1,\nu}, w^\nu)$ ... eminently parallelizable

PH: Implementation

implementation: choice of ρ ... scenario (×), decision (+) dependent

(heuristic) extension to problems with integer variables

non-convexities: e.g. ground-water remediation with non-linear PDE recourse

asynchronous

partitioning (= different information feeds)

$$\min \mathbb{E}\left\{ f(\boldsymbol{\xi},x) \right\}, \quad f(\xi,x) = f_0(x) + \iota_{C(\xi,x)}(x)$$

$S = \left\{ \Xi_1, \Xi_2, \ldots, \Xi_N \right\}$ a partitioning of Ξ, $\quad p_n = \mu(\Xi_n)$

$\mathbb{E}\left\{ f(\boldsymbol{\xi},x) \right\} = \sum_n p_n \mathbb{E}\left\{ f(\boldsymbol{\xi},x) \mid \Xi_n \right\}$ (Bayes)

defining $g(n,x) = \mathbb{E}\left\{ f_0(\boldsymbol{\xi},x) \mid \Xi_n \right\}$ if $x \in C_n = \bigcap_{\xi \in \Xi_n} C_\xi$

solve the problem as: $\min \sum_{n=1}^{N} p_n g(n,x)$

Bundling

Multistage Stochastic Programs

$\min_{x\in\mathcal{N}^a} \mathbb{E}\left\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\right\}, \quad x(\xi) = \left(x^1(\xi),\ldots,x^T(\xi)\right)$

filtration : $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_T = \mathcal{A}, \ \mathcal{A}_0$ trivial

$x \in \mathcal{N}^a$ if x^t $\mathcal{A}_{t-1}$-measurable $\approx \sigma$-field($\overset{\to\nu-1}{\boldsymbol{\xi}}$)

(here $\boldsymbol{\xi}^0$ deterministic, $x^1(\boldsymbol{\xi}) \equiv x^1$)

under usual $\mathbb{C}.\mathbb{Q}.$ (convex case): $\bar{x} \in \mathcal{X}$ optimal if

$\exists\, \bar{w} \perp \mathcal{N}^a, \bar{w} \in \mathcal{X}^*$ such that $\bar{x} \in \arg\min_{x\in\mathcal{X}} Ef(x) - \mathbb{E}\left\{\langle \bar{\boldsymbol{w}}, x\rangle\right\}$

$\bar{w} \perp \mathcal{N}^a \Leftrightarrow \mathbb{E}\left\{\bar{w}(\boldsymbol{\xi}) \middle| \mathcal{A}_{t-1}\right\} = 0, \forall t = 1,\ldots,T$

$\bar{w}$ non-anticipativity prices

at which to buy the right to adjust decision (after observation)

can be viewed as insurance premiums, ...

PH: binary variables

$\min\langle c, x\rangle + \sum_{\xi\in\Xi} p_\xi \langle q_\xi, y_\xi\rangle$ such that

$x \in C_1, \;\; y_\xi \in C_2(\xi, x) \;\forall\, \xi \in \Xi$

binary (integer) variables: some x's, some y_ξ's.

PH: binary variables

$\min\langle c, x\rangle + \sum_{\xi\in\Xi} p_\xi \langle q_\xi, y_\xi\rangle$ such that

$x \in C_1, \;\; y_\xi \in C_2(\xi, x) \; \forall\, \xi \in \Xi$

binary (integer) variables: some x's, some y_ξ's.

Choice of $\rho \to \rho_j$ depending on $c_j, |x_j|, \ldots$ and augmentation

Variable Fixing, in particular binaries, $x_j(s) =$ constant (k iterations)
Variable Slamming: aggressive variable fixing $x_j(s) \approx$ constant (& $c_j x_j(s)$)
"Sufficient" variable convergence $\sim$ for small values of $c_j x_j(s)$

Termination criterion: variable slamming when $x_j^\nu(\xi) - x_j^{\nu+1}(\xi)$ small

Detecting cycling behavior: (simple) hashing scheme

PH: binary variables

$\min\langle c, x\rangle + \sum_{\xi\in\Xi} p_\xi \langle q_\xi, y_\xi\rangle$ such that

$x \in C_1, \;\; y_\xi \in C_2(\xi, x)\; \forall\, \xi \in \Xi$

binary (integer) variables: some x's, some y_ξ's.

Choice of $\rho \to \rho_j$ depending on $c_j, |x_j|, \ldots$ and augmentation

Variable Fixing, in particular binaries, $x_j(s) =$ constant (k iterations)

Variable Slamming: aggressive variable fixing $x_j(s) \approx$ constant (& $c_j x_j(s)$)

"Sufficient" variable convergence $\sim$ for small values of $c_j x_j(s)$

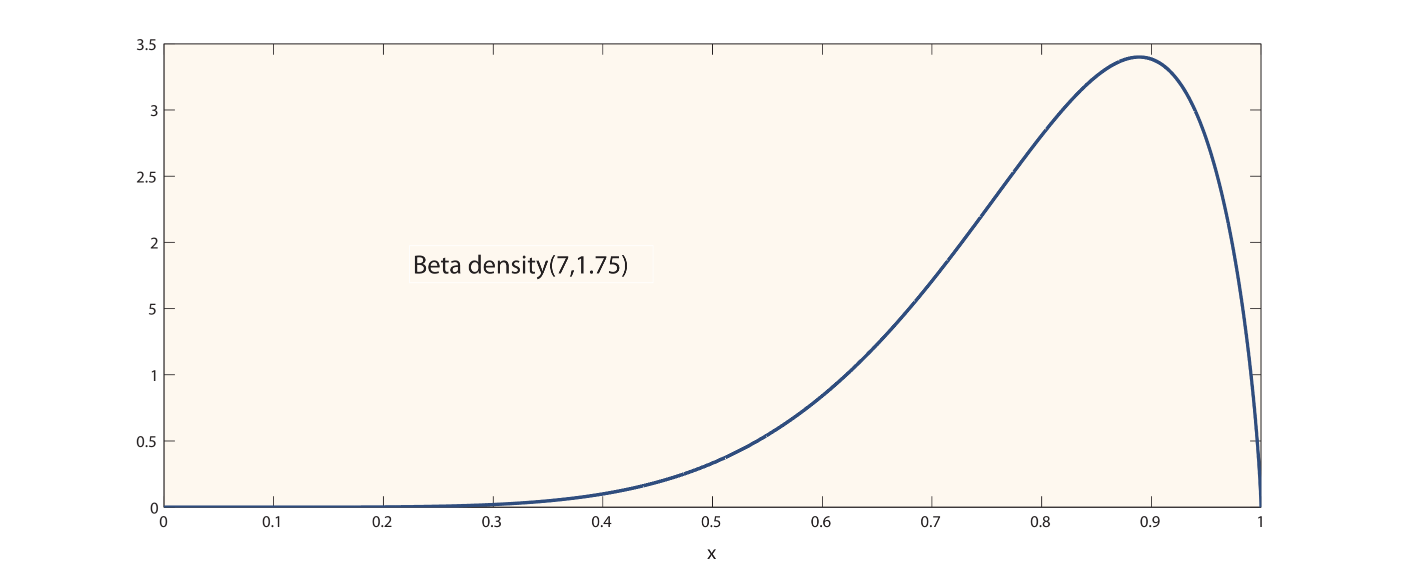
Termination criterion: variable slamming when $x_j^\nu(\xi) - x_j^{\nu+1}(\xi)$ small

Detecting cycling behavior: (simple) hashing scheme

Enough variables fixed ⇒ clean up with CPLEX-MIP

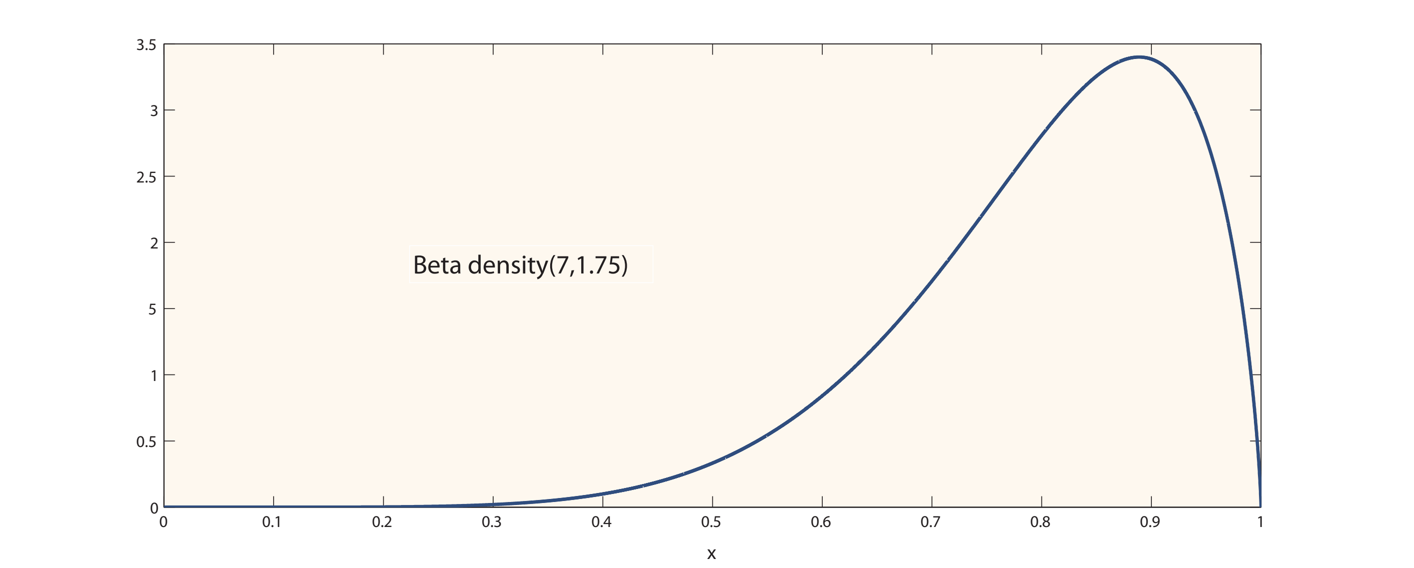
Augmentation function

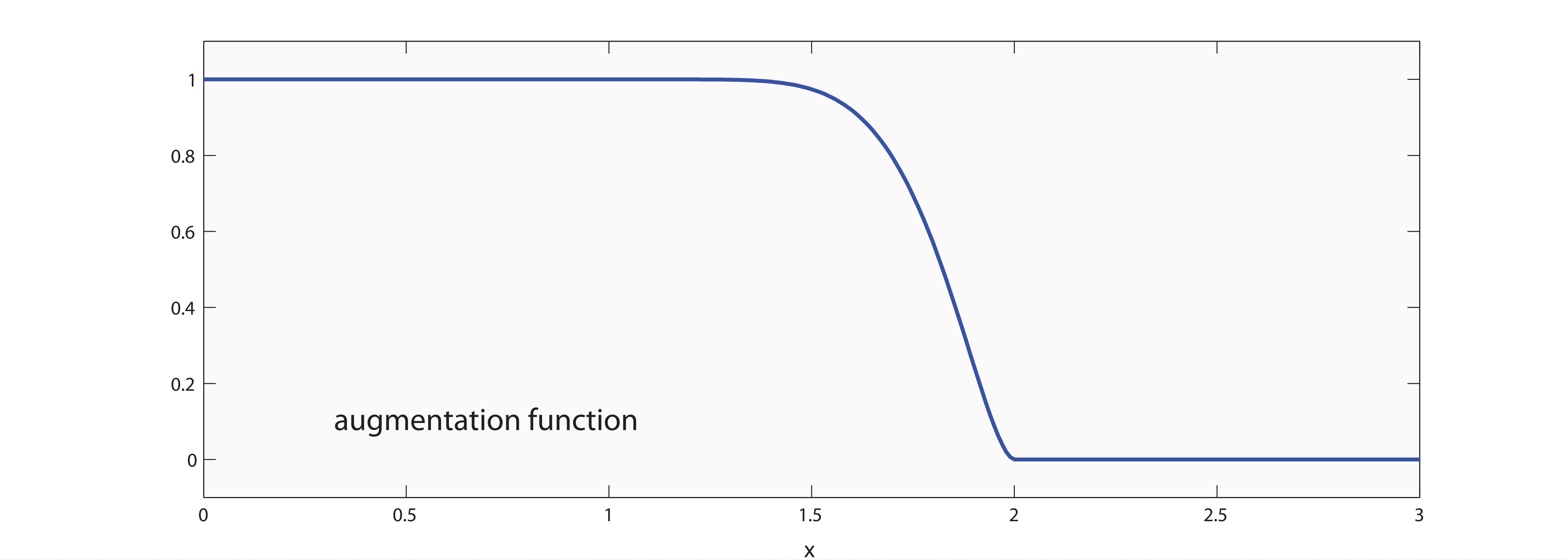
$$m(\Delta, \lambda_{\max}; z, \lambda) = \int_0^{\Delta} \psi(z - s, \lambda, \lambda_{\max}) \varphi(\Delta; s)\, ds, \quad z \in [0, \lambda_{\max}]$$



Augmentation function

$$m(\Delta, \lambda_{\max}; z, \lambda) = \int_0^{\Delta} \psi(z - s, \lambda, \lambda_{\max}) \varphi(\Delta; s)\, ds, \quad z \in [0, \lambda_{\max}]$$





Error Bounds

$f(\xi,x) = f_0(\xi,x^1,x^2)$ if $x^1 \in C^1, x^2 \in C^2(\xi,x^1)$; $+\infty$ else

Stochastic Program (P):

$$\min_{x \in \mathcal{M}} \mathbb{E}\left\{f(\xi,x)\right\} \text{ such that } x^1 \equiv \mathbb{E}\{x^1_\xi\}$$

Dual Program (D)

$$\max_{w \in \mathcal{M}^*} \mathbb{E}\left\{-f^*(\xi,w_\xi)\right\} \text{ such that } \mathbb{E}\{w_\xi\} = 0$$

weak duality holds: $\inf P \geq \sup \mathrm{D} \Rightarrow$ for any feasible $\hat{w}$

$$-f^*(\xi,\widehat{w}_\xi) = \min_x \left[f(\xi,x) + \left\langle \widehat{w}_\xi, x^1 \right\rangle, x \in \mathbb{R}^{n_1+n_2} \right]$$

yields a lower bound for (P), better if $\widehat{w}_\xi$ is near-optimal

$\Rightarrow$ rely on w^* of PH-algorithm to generate lower bound.

Unit Commitment SCUC
(PH with binary variables)

ARPA-e Project

Sandia National Labs, Iowa State Univ.,
Univ. of California-Davis, Alstom, New-England ISO

Transmission Network

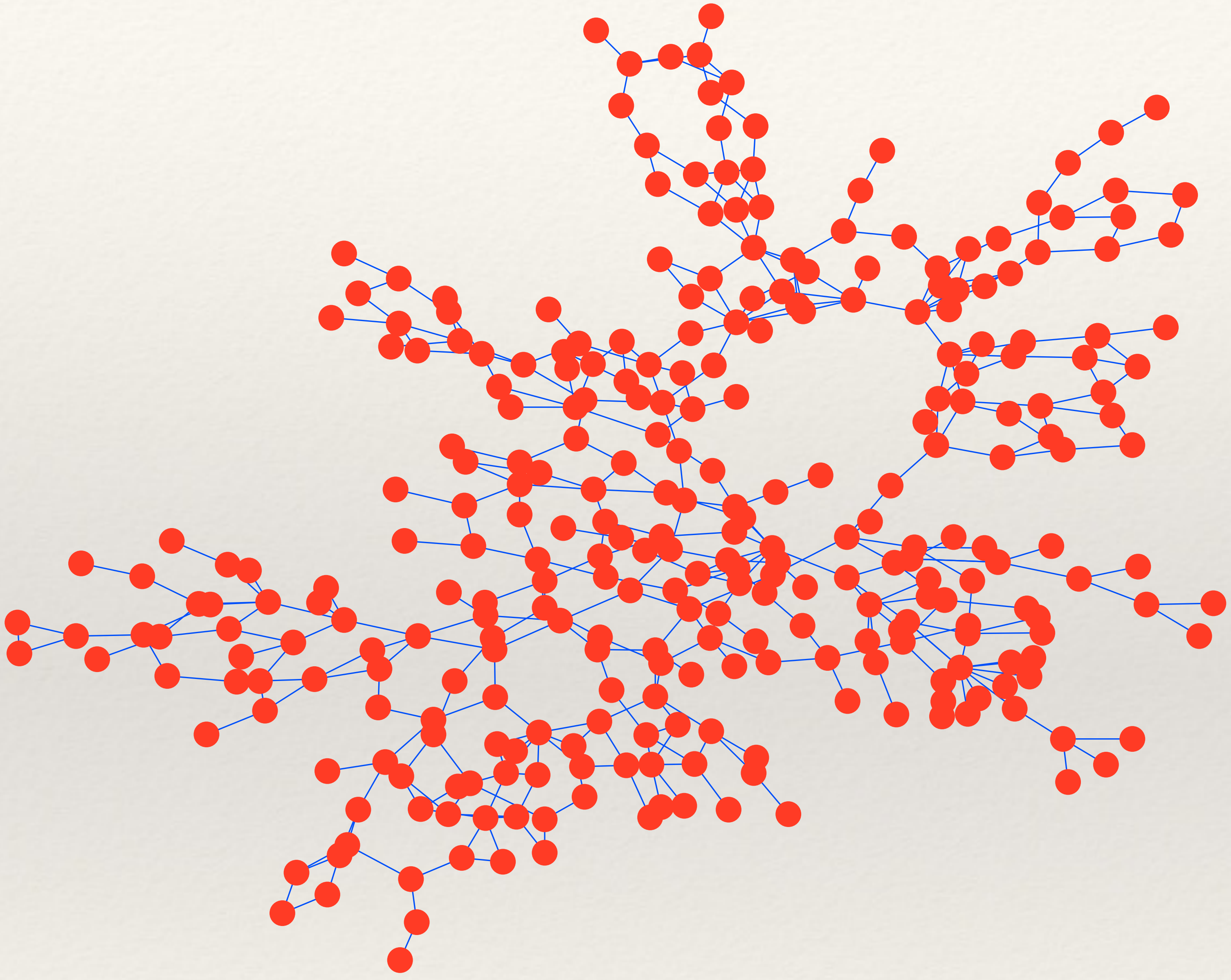
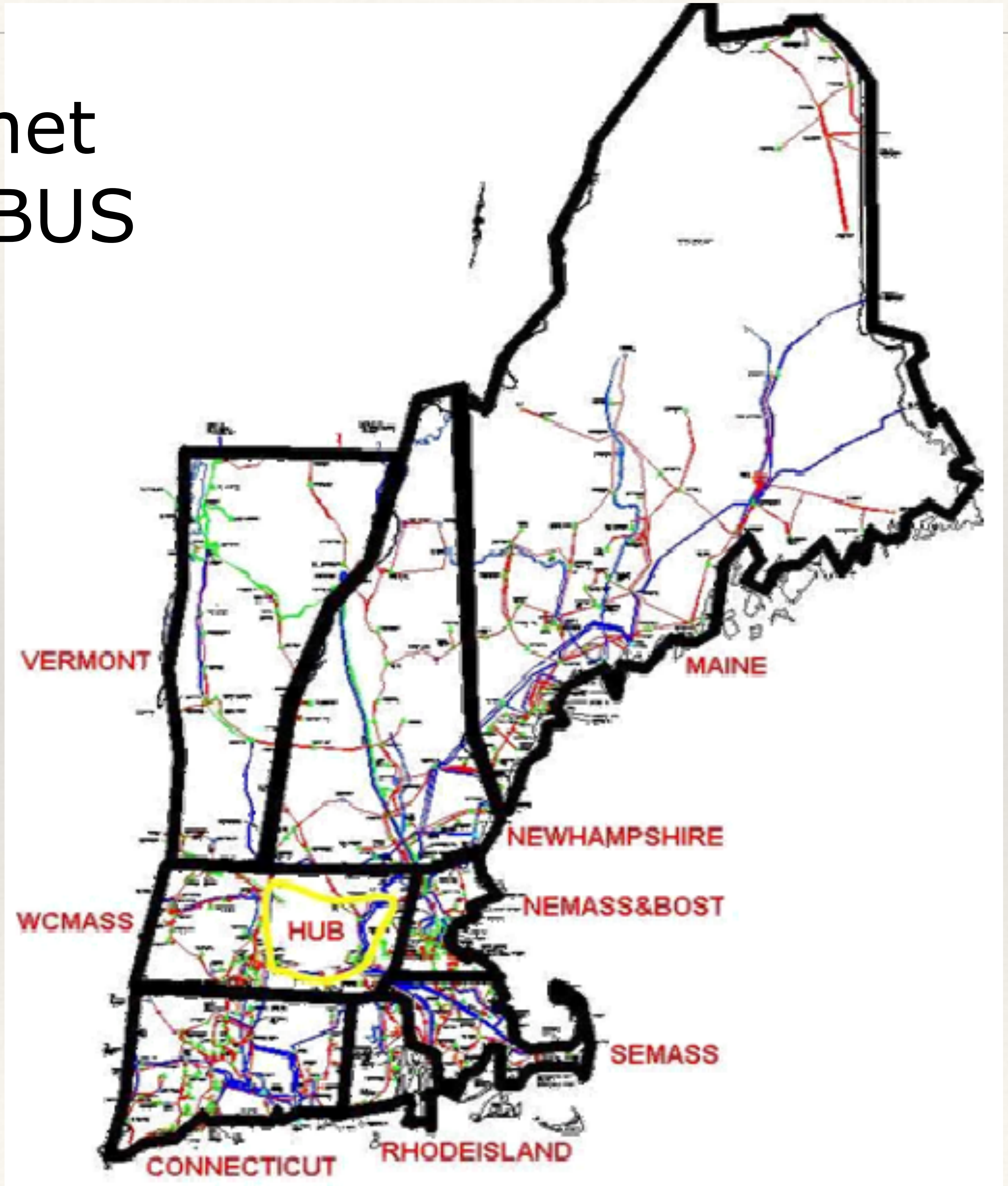


Figure 1. Topology of the IEEE 300 node system

Transmission Network

NE-ISO net
~30,000 BUS



ISO: Independent System Operator

FERC

Federal Energy Regulatory Commission

↓

RTO

In the US is an organization that is responsible for moving electricity over large interstate areas; coordinates, controls and monitors an electricity transmission grid that is larger with much higher voltages than the typical power company's distribution grid.

↓

ISO

*Is an organization formed at the direction or recommendation of the **FERC**, in the areas where an ISO is established, it coordinates, controls and monitors the operation of the electrical power system, usually within a single US State, but sometimes encompassing multiple states.*

ISO New England Inc. (*ISO-NE*) *is an independent, non-profit RTO, serving Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont. Its Board of Directors and its over 400 employees* have no financial interest or ties to any company doing business in the region's wholesale electricity marketplace.

Energy Sources

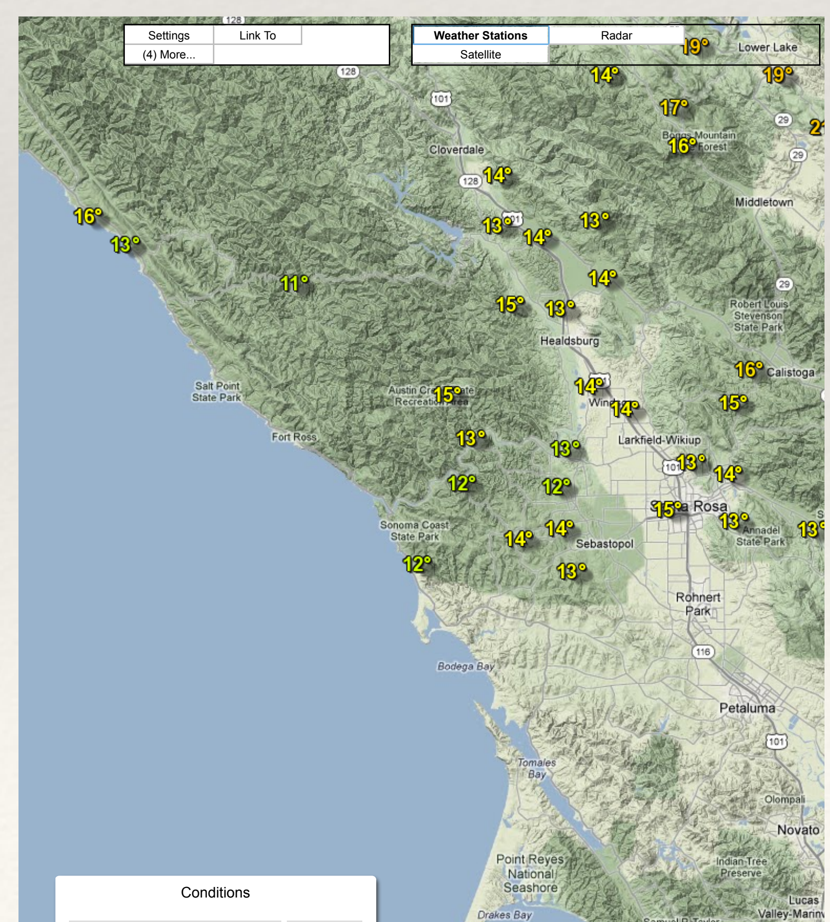
- nuclear energy
- hydro-power
- thermal plants (coal, oil, shale oil, bio, rubish, . . .)
- gas turbines (natural gas, from ”cracking’)
- renewables (wind, solar, ..., ocean waves)

different characteristics

Uncertainties

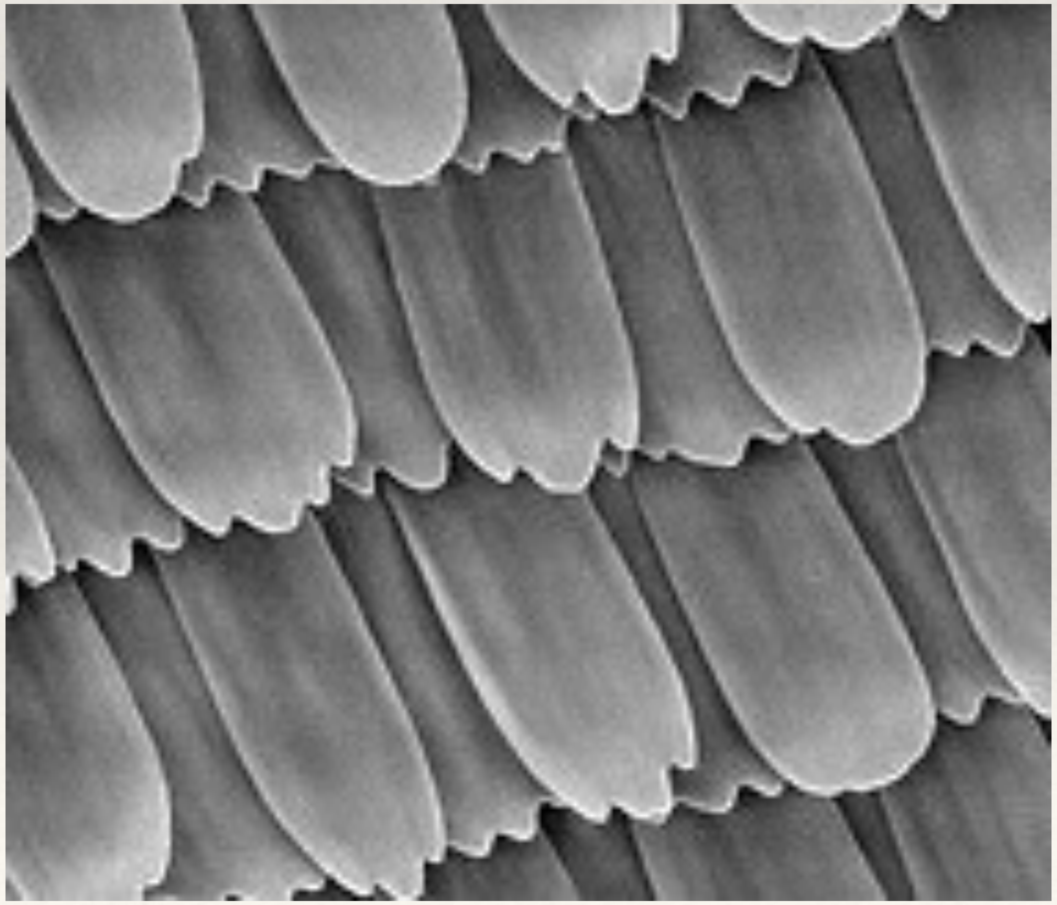
- WEATHER: demand & supply (especially renewables)
- industrial-commercial environment (demand)
- seasonal, day of the week, time of the day
- contingencies: transmission lines, generators

El Cerrito, California (94530) Conditions & Forecast http://www.wunderground.com/cgi-bin/findweather/getForecast...



Uncertainties

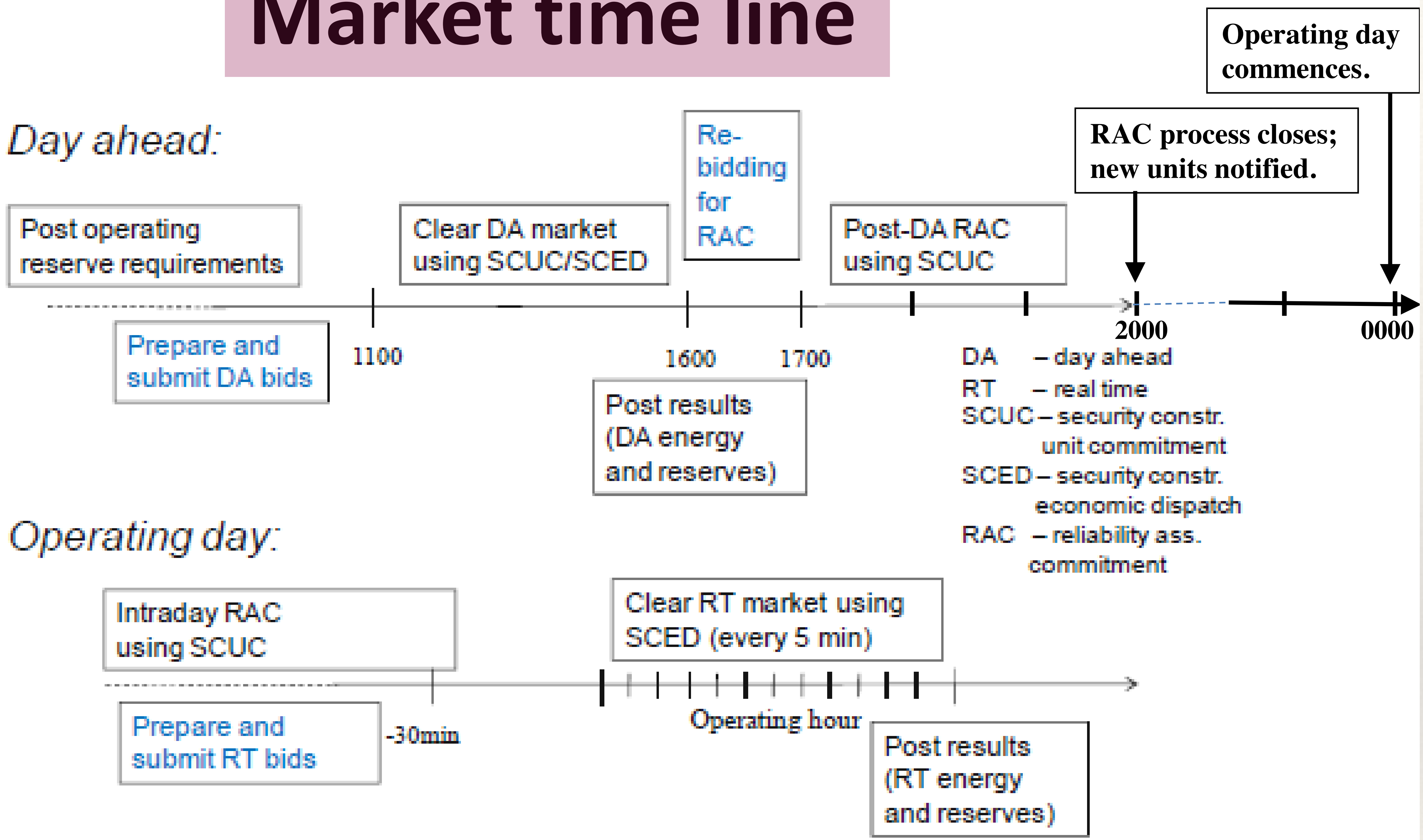
- WEATHER: demand & supply (especially renewables)
- industrial-commercial environment (demand)
- seasonal, day of the week, time of the day
- contingencies: transmission lines, generators



Uncertainties

- WEATHER: demand & supply (especially renewables)
- industrial-commercial environment (demand)
- seasonal, day of the week, time of the day
- contingencies: transmission lines, generators

Market time line



	MISO	NYISO	PJM	ERCOT	CAISO
Market timeline	DA offers due: 11am DA results: 4pm Re-bidding due: 5pm RT offers due: OH -30 min	DA offers due: 5 am DA results: 11 am RT offers due: OH -75 min	DA offers due: noon DA results: 4pm RT offers due: 6pm DA	DA bids due (reserves): 1pm/4pm DA results (reserves): 1.30pm/6pm RT offers due: OH -60 min	DA offers: 10am DA results: 1pm RT offers: OH - 75 min

Ref: A. Botterud, J. Wang, C. Monteiro, and V. Miranda "Wind Power Forecasting and Electricity Market Operations," available at www.usaee.org/usaee2009/submissions/OnlineProceedings/Botterud_etal_paper.pdf

Short history of ISO-management techniques

- RT: deterministic optimization with LMP (dual variables associated with demand(s) constraints).
- SCUC/SCED: Lagrangian relaxation with conservative reliability constraints
- SCUC/SCED: deterministic MIP with conservative RUT
- ARPA-”E (project): “take into account uncertainty”

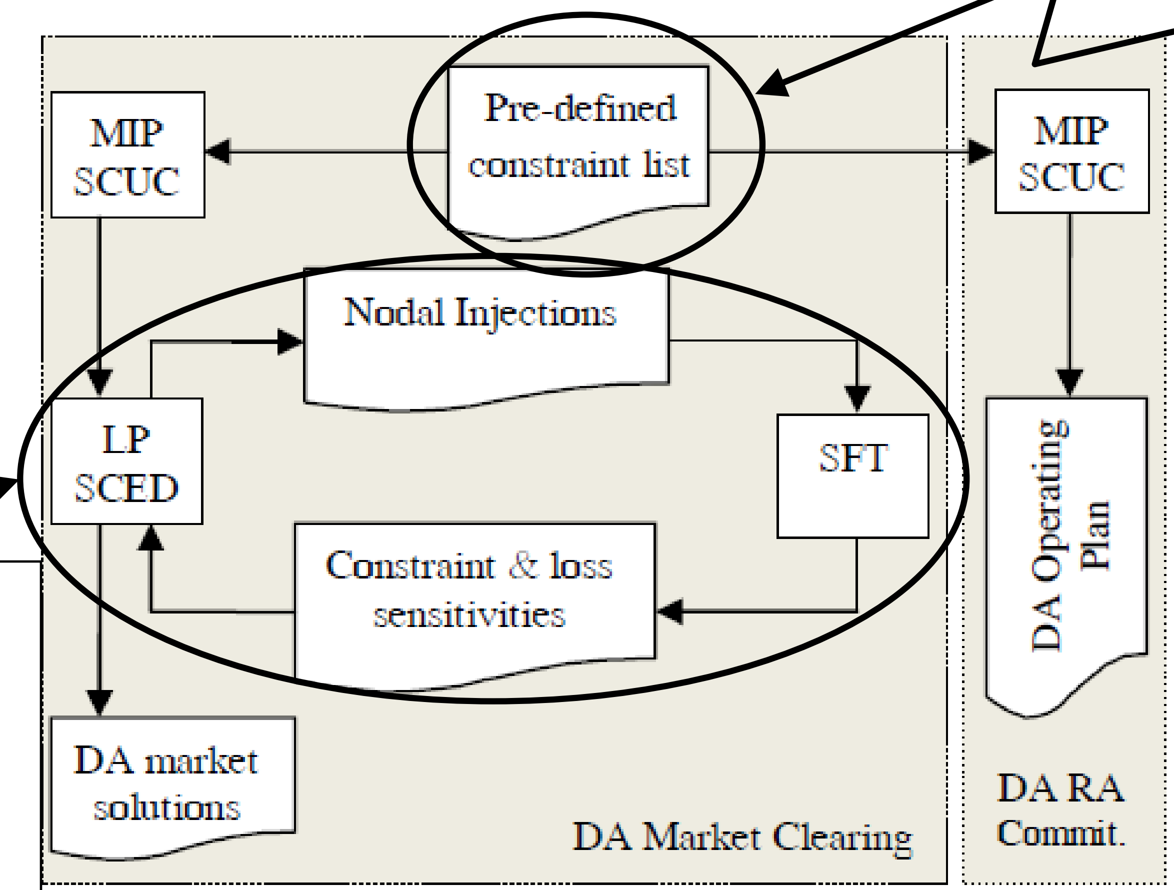
A collection of stochastic-programs

- DA-SCUC/SCED unit commitment *binaries*
- DA-RAC rebidding assessment bidding *(binaries)*
- DA-RUT - reliability commitments (spinning, N-1)
- RT - 3 min (real time adjustments) LMP's
- SCED2 - 3 or 4 hours schedule to foresee ramp ups/down, etc.

DA = day ahead

Day-Ahead Market

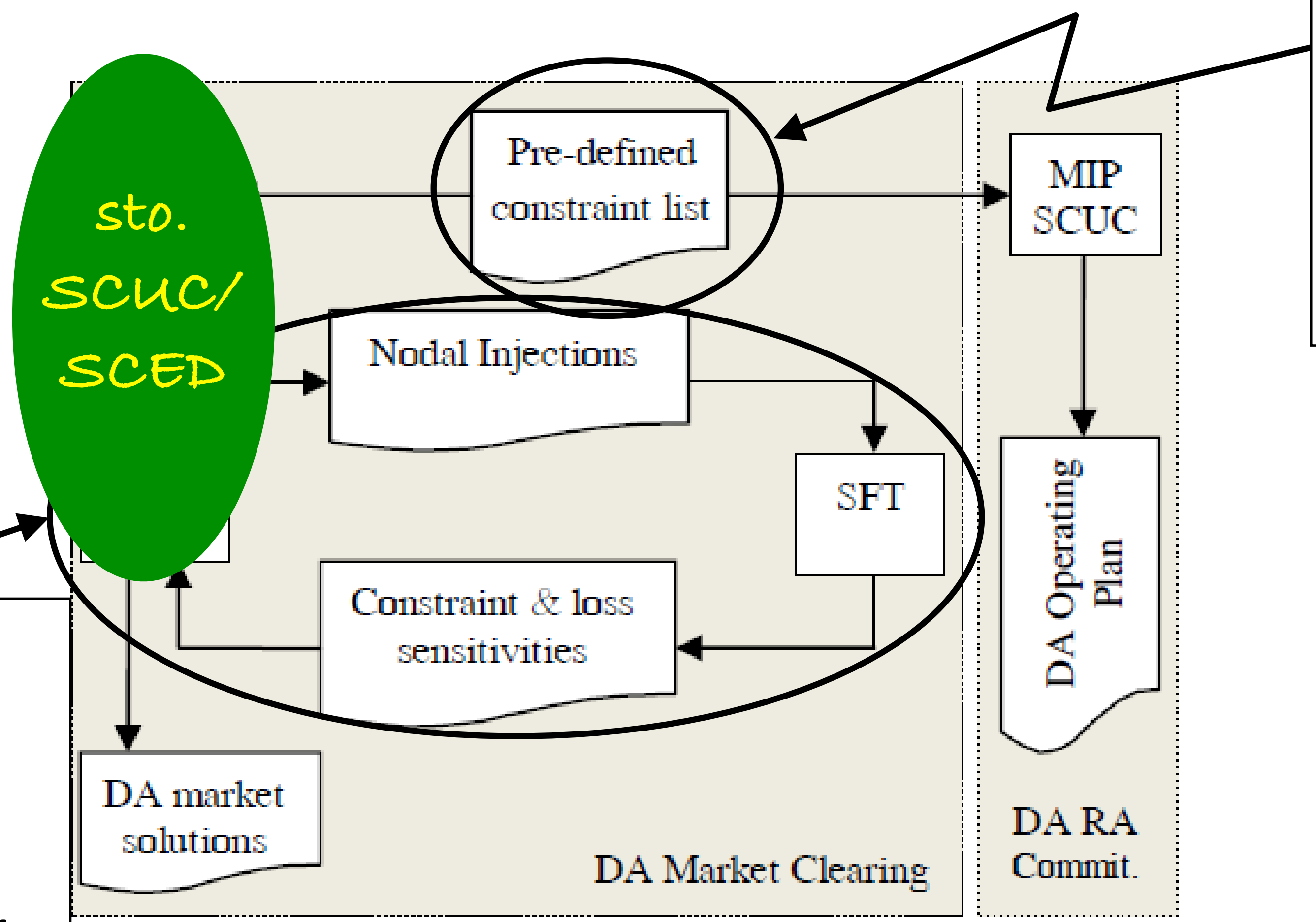
SCUC enforces limited number of transmission constraints on the commitment solution.



Each hourly SCED performs SFT, which tests all contingencies in a list and for violations, imposes appropriate constraints in SCED and resolves it.

Ref: Xingwang Ma, Haili Song, Mingguo Hong, Jie Wan, Yonghong Chen, Eugene Zak, “The Security-constrained Commitment and Dispatch For Midwest ISO Day-ahead Co-optimized Energy and Ancillary Service Market,” Proc. of the 2009 IEEE PES General Meeting.

Day-Ahead Market

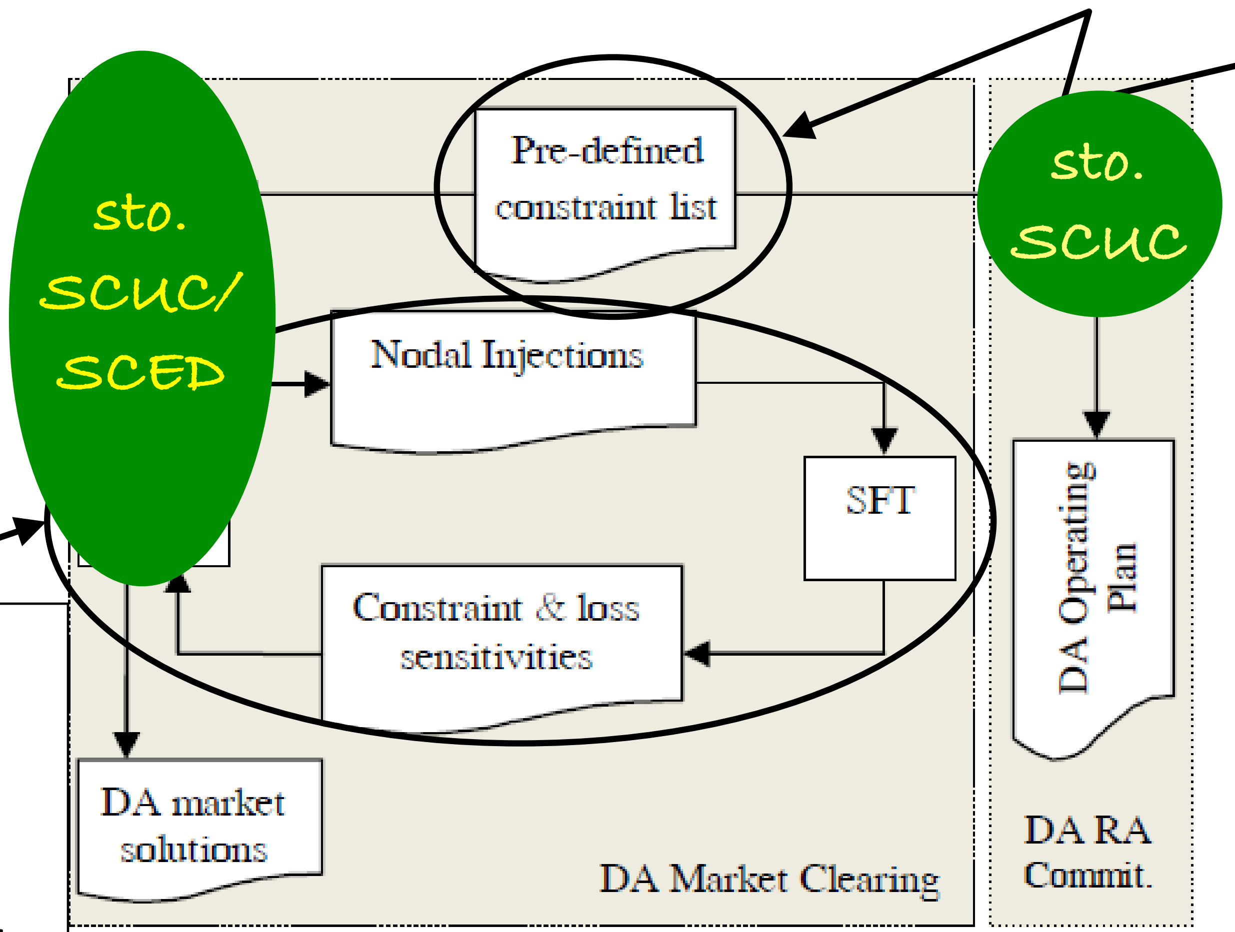


SCUC enforces limited number of transmission constraints on the commitment solution.

Each hourly SCED performs SFT, which tests all contingencies in a list and for violations, imposes appropriate constraints in SCED and resolves it.

Ref: Xingwang Ma, Haili Song, Mingguo Hong, Jie Wan, Yonghong Chen, Eugene Zak, "The Security-constrained Commitment and Dispatch For Midwest ISO Day-ahead Co-optimized Energy and Ancillary Service Market," Proc. of the 2009 IEEE PES General Meeting.

Day-Ahead Market

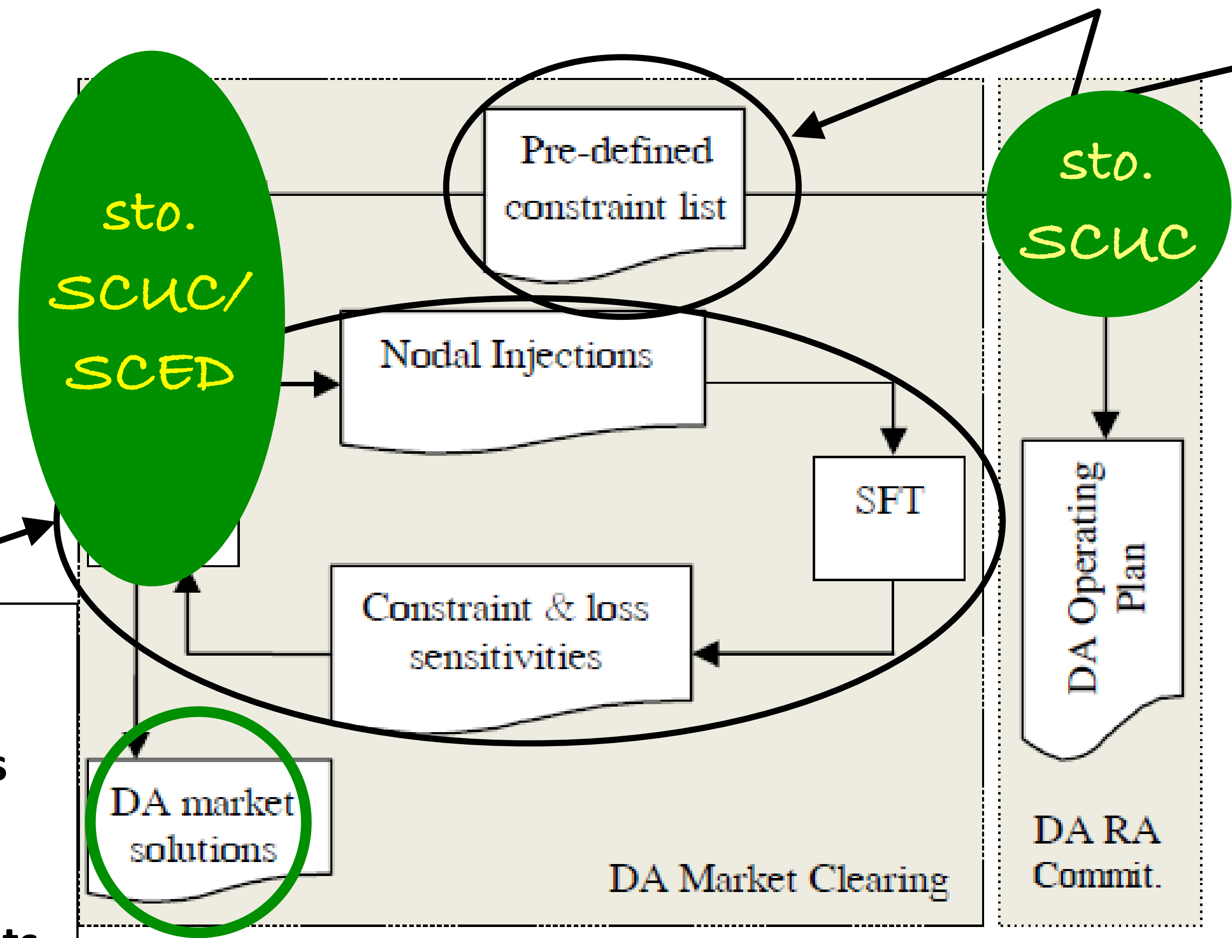


SCUC enforces limited number of transmission constraints on the commitment solution.

Each hourly SCED performs SFT, which tests all contingencies in a list and for violations, imposes appropriate constraints in SCED and resolves it.

Ref: Xingwang Ma, Haili Song, Mingguo Hong, Jie Wan, Yonghong Chen, Eugene Zak, "The Security-constrained Commitment and Dispatch For Midwest ISO Day-ahead Co-optimized Energy and Ancillary Service Market," Proc. of the 2009 IEEE PES General Meeting.

Day-Ahead Market



SCUC enforces limited number of transmission constraints on the commitment solution.

Each hourly SCED performs SFT, which tests all contingencies in a list and for violations, imposes appropriate constraints in SCED and resolves it.

Ref: Xingwang Ma, Haili Song, Mingguo Hong, Jie Wan, Yonghong Chen, Eugene Zak, “The Security-constrained Commitment and Dispatch For Midwest ISO Day-ahead Co-optimized Energy and Ancillary Service Market,” Proc. of the 2009 IEEE PES General Meeting.

Abstract Unit Commitment

min. expectation
(actually: risk measure)
with penalties

production cost startup cost shutdown cost

$$\text{Minimize} \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \quad \text{with}$$

K time periods J generating units

power output, demand

$$\sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K$$

adjust node balance eq'ns

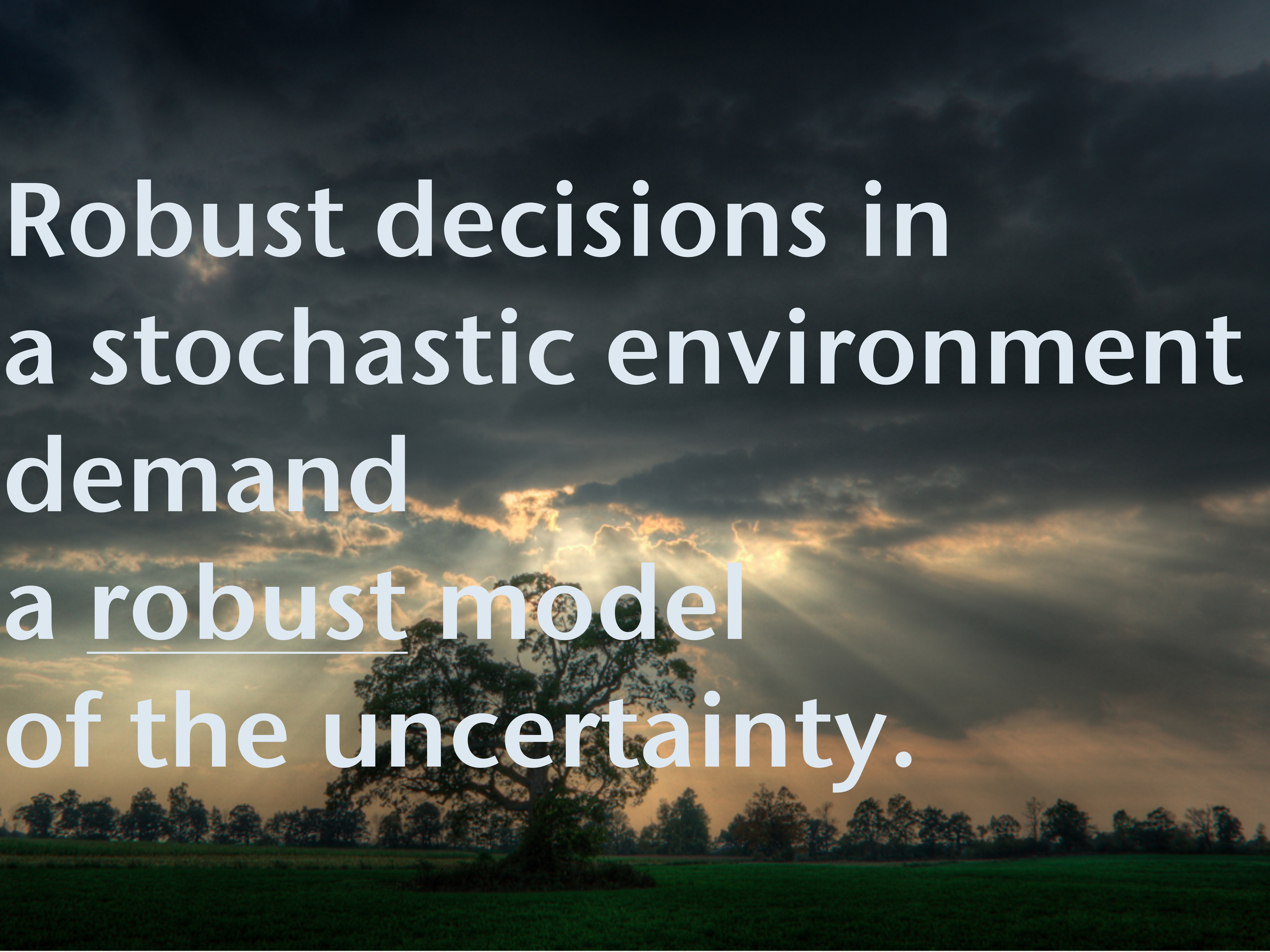
max power output, spinning reserve

$$\sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K$$

$$p_j(k), \overline{p}_j(k) \in \Pi, \quad \forall j \in J, \; \forall k \in K$$

Π region of feasible production, all generating units, all time periods.
The specific nature of Π is model-dependent.

"Stochastic Version"


Robust decisions in
a stochastic environment
demand
a robust model
of the uncertainty.

Solution procedures

$$\min_{x\in\mathcal{N}^a} \mathbb{E}\left\{f(\xi,x(\xi))\right\} = \min_{x^1\in\mathbb{R}^{n_1}} f_1(x^1) + EQ_1(x^1)$$

$$EQ_1(\xi;x^1) = \mathbb{E}\left\{\inf_{x^2\in\mathbb{R}^{n_2}} f_2(\xi;x^1,x^2) + EQ_2(\xi;x^1,x^2)\middle|\mathcal{A}_1\right\}$$

$$EQ_2(\xi;x^1,x^2(\xi)) = \mathbb{E}\left\{\inf_{x^3\in\mathbb{R}^{n_3}} f_3(\xi;x^1,x^2(\xi),x^3)\middle|\mathcal{A}_2\right\}$$

deterministic optimization! convex when f convex random lsc function

in theory: any algorithmic procedure!

hurdles: values, (sub)gradients, "Hessians" of $f_1(x^1) + EQ_1(x^1)$

are either not acessible or at best, prohibitively EXPENSIVE

Approaches: $P^\nu \sim P \Rightarrow$ approximating stochastic process $\left\{\xi_t, t \le T\right\}$

sampling: a) same as approximation except P^s random measure

b) SAA-strategy for $\partial\left(\mathbb{E}\left\{f(\xi,x(\xi))\right\} + N_{\mathcal{N}^a}(x(\xi))\right)$

Deterministic Equivalent

$$\min_{x\in\mathcal{N}^a} \mathbb{E}\{f(\xi,x(\xi))\} = \mathbb{E}\left\{\mathbb{E}\cdots\left\{\mathbb{E}\left\{f(\xi,x(\xi))\middle|\mathcal{A}_T\right|\cdots\middle|\mathcal{A}_1\middle|\mathcal{A}_0\right\}\right\}\right\}$$

"time-staged objective":

$$= f_1(x^1) + \mathbb{E}\left\{f_2(\xi;x^1,x^2(\xi)) + \mathbb{E}\left\{f_3(\xi;x^1,x^2(\xi),x^3(\xi)\middle|\mathcal{A}_2\right\}\middle|\mathcal{A}_1\right\} \quad \cdots$$

$$= f_1(x^1) + \mathbb{E}\left\{f_2(\xi;x^1,x^2(\xi)) + EQ_2(\xi;x^1,x^2(\xi))\middle|\mathcal{A}_1\right\}$$

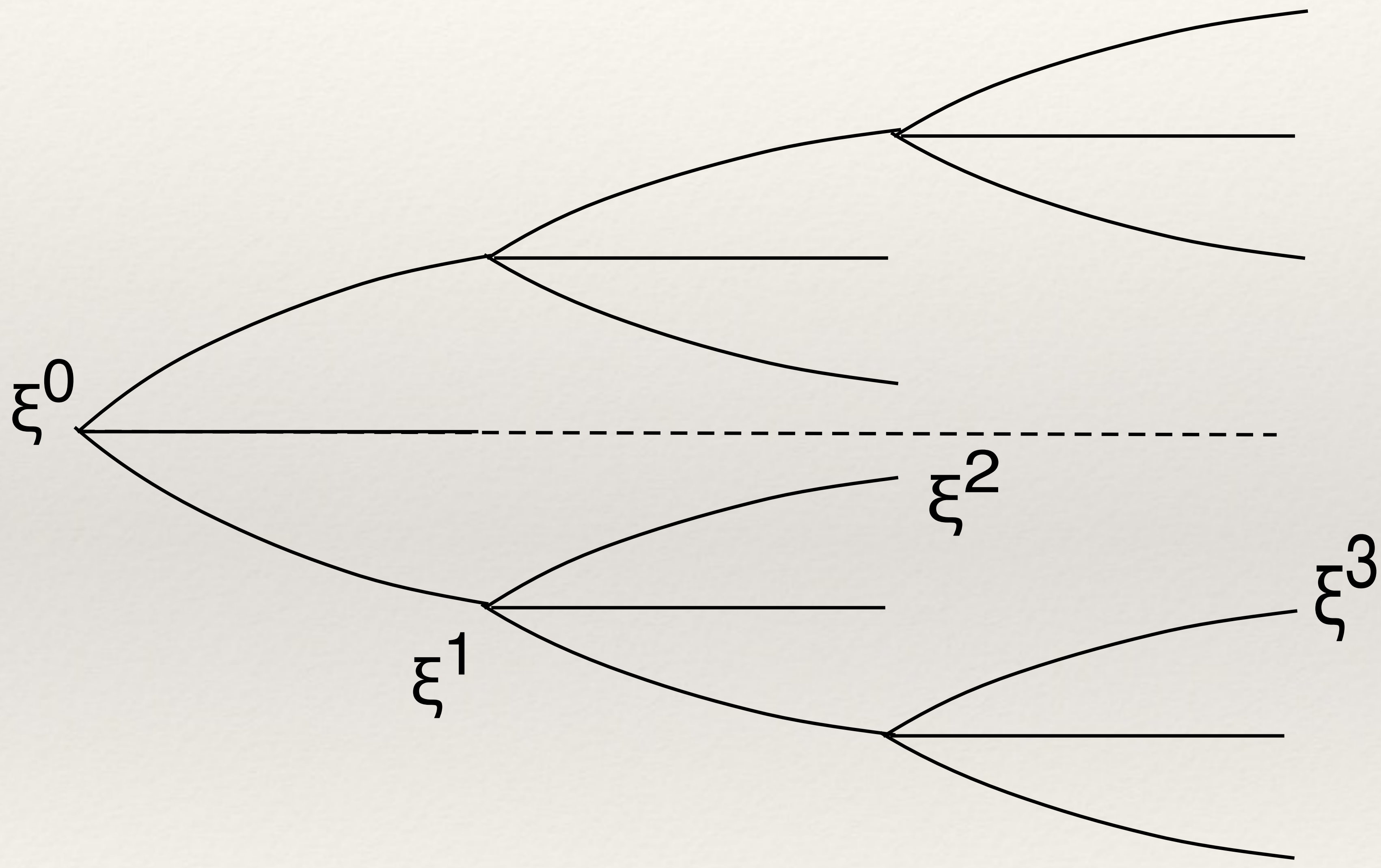
$$EQ_2(\xi;x^1,x^2(\xi)) = \mathbb{E}\left\{\inf_{x^3\in\mathbb{R}^{n_3}} f_3(\xi;x^1,x^2(\xi),x^3)\middle|\mathcal{A}_2\right\}$$

$$= f_1(x^1) + \mathbb{E}\left\{EQ_1(\xi;x^1,x)\middle|\mathcal{A}_1\right\}$$

$$EQ_1(\xi;x^1) = \mathbb{E}\left\{\inf_{x^2\in\mathbb{R}^{n_2}} f_2(\xi;x^1,x^2) + EQ_2(\xi;x^1,x^2)\middle|\mathcal{A}_1\right\}$$

$$= f_1(x^1) + EQ_1(x^1)$$

Discrete Scenario Tree



Sequential l.p. Strategy

$\min f_0(x),\;\; x \in X \in \mathbb{R}^n,\quad f_0$ linear (not essential)

$f_i(x) \le 0,\;\; i = 1,\dots,s,\;\; f_i(s) = 0, i = s+1,\dots,m$ (affine)

in the $s+1$ first constraints: $f_i(x) = \sup_{t\in T} f_{i,t}(x),\;\; f_i \ge f_{i,t}$ affine

0. $\nu = 0$, pick polytope (box) $K^0 \ni x^{opt}$

1. $x^\nu \in \arg\min f_0$ on K^ν, set $\mathrm{i}_\nu : f_{i_\nu}(x^\nu) = \max_{1\le i\le s} f_i(x^\nu)$

   if $f_{i_\nu}(x^\nu) \le 0, x^\nu$ optimal, otherwise go to 2.

2. return to 1. with $K^{\nu+1} = K^\nu \cap \left\{ \left\langle \nabla f_{i_\nu}(x^\nu), x - x^\nu \right\rangle + f_{i_\nu}(x^\nu) \le 0 \right\}$

when f_0 is not linear (but convex): $\min \theta$ such that $f_0(x) - \theta \le 0$

convergence: finite # of steps or iterates cluster to optimal sol'n

SLP for Stochastic Programs

$\min f_1(x) + EQ_1(x)$ s.t. $Ax = b,\ x \geq 0 \quad (x = x^1)$

$$EQ_1(x) = \sum_{l=1}^{L} p_l Q_1(\xi^l, x) \quad L \text{ large}$$

$$Q_1(\xi^l, x) = \inf_{x^2 \in X_2} \left\{ f_2(\xi^l; x, x^2) + (EQ_2(\cdots) \right\}$$

$$\text{dom } EQ_1 = \bigcap_{l=1}^{L} \text{dom } Q_1(\xi^l, \cdot) = \bigcap_{l=1}^{L} \left\{ x \,\middle|\, \exists x^2 \in X_2, f_2(\xi^l; x, x^2) < \infty \right\}$$

0. $\nu = r = s = 0$

1. $\nu = \nu + 1$, solve: $\min f_1(x) + \theta,\ Ax = b,\ x \geq 0$ such that

(feasibility cuts) $\langle E_k, x \rangle \geq e_k,\ k = 1 \to r$

(optimality cuts) $\langle F_k, x \rangle + \theta \geq f_k,\ k = 1 \to s$

2. generate feasibility cuts: check if $x \in \text{dom } EQ_1$.

   No: E_k separates x from dom EQ_1, go to 1. Yes, go to 3.

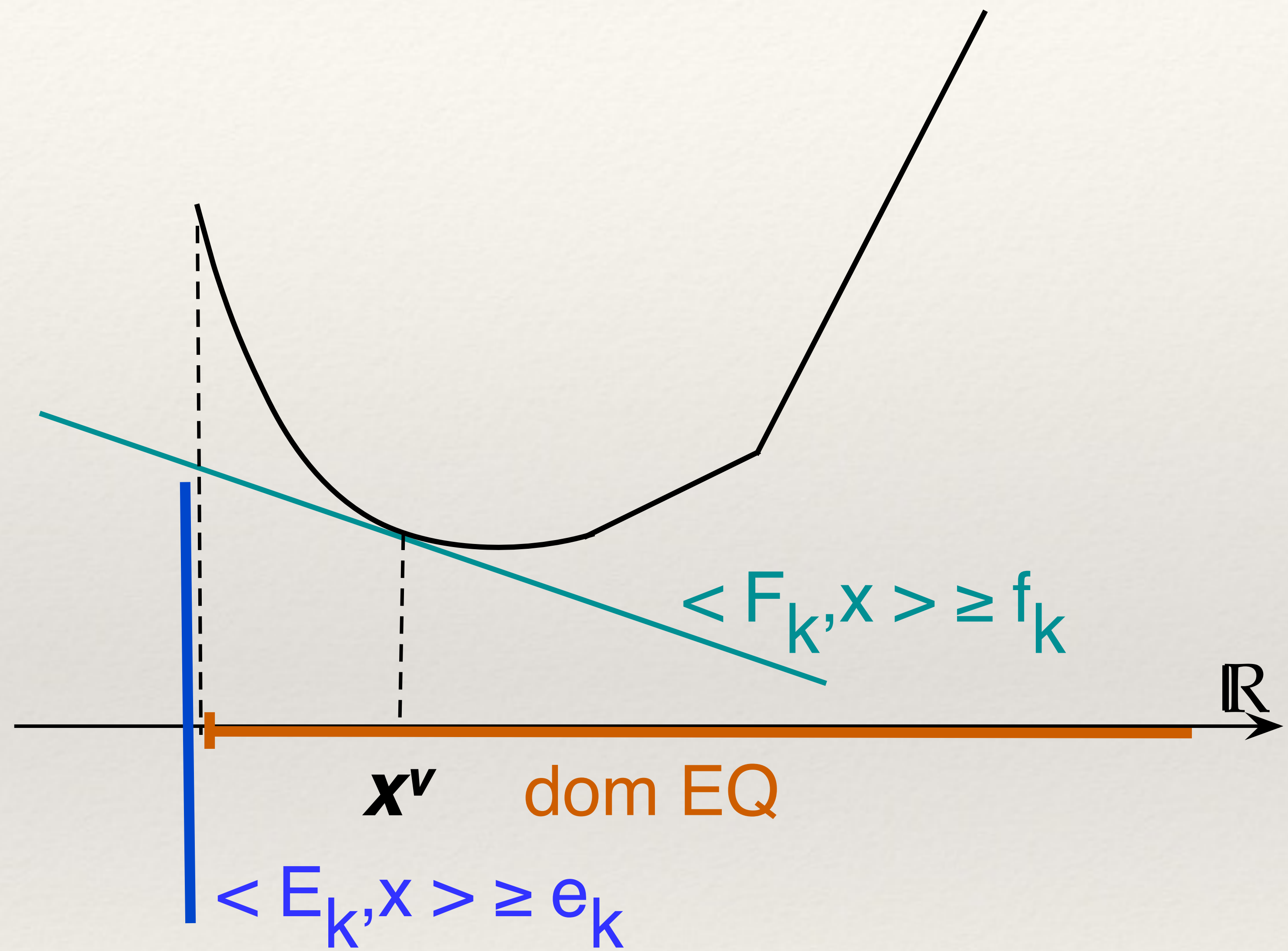
3. generate optimality cuts: $F_k \in \partial EQ_1(x^k)$, go to 1.

Generating cutting Hyperplanes

Generating cutting Hyperplanes

x^v

Generating cutting Hyperplanes



just a bit of “math”

Expectation Functionals

Expectation of $\overline{\mathbb{R}}$-valued functions (Fatou, monotone convergence, ...):

$$E\{f(\boldsymbol{\xi})\} = \int_\Xi f(\xi)P(d\xi) = \begin{cases} \infty & \text{if } P([f(\boldsymbol{\xi}) = \infty]) > 0 \\ \int_\Xi f(\xi)P(d\xi) & \text{otherwise,} \end{cases}$$

or $E\{f(\boldsymbol{\xi})\} = E\{\max[f(\boldsymbol{\xi}), 0]\} - E\{\max[-f(\boldsymbol{\xi}), 0]\}, \ \ \infty - \infty = \infty$ (convention).

.

$f : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}}, \quad Ef : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, assume $Ef \not\equiv \infty$

- **Convexity**. $x \mapsto f(\xi, x)$ is convex (resp. affine, sublinear), then so is Ef.
- **Lower semicontinous**. $x \mapsto f(\xi, x)$ lsc & convex or summably bounded below $\Rightarrow Ef$ lsc.
- **Subdifferentials**. Ef finite near x, for all $\xi \in \Xi$, $f(\xi, \cdot)$ convex, then

$$\partial Ef(x) = \mathbb{E}\{\partial f(\boldsymbol{\xi}, x)\} = \left\{ \int_\Xi v(\xi)\, P(d\xi) \;\middle|\; v \text{ integrable}, v(\xi) \in \partial f(\xi, x) \right\}.$$

.

Characterization of minimizers

Theorem. Ef an expectation functional with $f(\xi,\cdot)$ convex.
Then, $x^0 \in \operatorname{argmin} Ef \iff \exists\, v : \Xi \to \mathbb{R}$, $\mathbb{E}\{v(\boldsymbol{\xi})\} = 0$, $v(\xi) \in \partial f(\xi, x^0)$, i.e.,

$$x^0 \in \operatorname*{argmin}_{x\in\mathbb{R}} \left\{ f(\xi,x) - v(\xi)x \right\} \quad \forall \xi \in \Xi$$

Proof. If $v(\cdot)$ exists, then $0 \in \partial Ef(x^0)$, i.e., $x^0 \in \operatorname{argmin} Ef$.

On the other hand, if $0 \in \partial Ef(x^0)$, $\exists\, v$ such that $\mathbb{E}\{v(\boldsymbol{\xi})\} = 0$ and $v(\xi) \in \partial f(\xi, x^0)$ is guaranteed by 'Subdufferential property'. The equivalence

$$v(\xi) \in \partial f(\xi,x^0) \;\&\; x^0 \in \operatorname{argmin}_x\{f(\xi,x) - v(\xi)x\}$$

is validated by Fermat's rule. $\square$

Characterization of minimizers

Theorem. Ef an expectation functional with $f(\xi, \cdot)$ convex.
Then, $x^0 \in \operatorname{argmin} Ef \iff \exists\, v : \Xi \to \mathbb{R}$, $\mathbb{E}\{v(\boldsymbol{\xi})\} = 0$, $v(\xi) \in \partial f(\xi, x^0)$, i.e.,

$$x^0 \in \operatorname*{argmin}_{x \in \mathbb{R}} \left\{ f(\xi, x) - v(\xi)x \right\} \quad \forall \xi \in \Xi$$

Proof. If $v(\cdot)$ exists, then $0 \in \partial Ef(x^0)$, i.e., $x^0 \in \operatorname{argmin} Ef$.

On the other hand, if $0 \in \partial Ef(x^0)$, $\exists\, v$ such that $\mathbb{E}\{v(\boldsymbol{\xi})\} = 0$ and $v(\xi) \in \partial f(\xi, x^0)$ is guaranteed by 'Subdufferential property'. The equivalence

$$v(\xi) \in \partial f(\xi, x^0) \;\&\; x^0 \in \operatorname{argmin}_x \{f(\xi, x) - v(\xi)x\}$$

is validated by Fermat's rule. $\square$

Knowing v allows the interchange of minimization and expectation